

De ontwikkeling van spraakmotorische controle I: Modellen

Cathelijne M. J. Y. Tesink^{1,2} en Ben Maassen³

¹ *Afdeling Psychiatrie, Universitair Medisch Centrum St. Radboud, Nijmegen*

² *F.C. Donders Centre for Cognitive Neuroimaging, Katholieke Universiteit Nijmegen*

³ *Interdisciplinair Kinderneurologisch Centrum/Afdeling Keel, Neus, Oorheelkunde/ Medische psychologie, Universitair Medisch Centrum St. Radboud, Nijmegen*

In dit literatuuroverzicht, dat bestaat uit twee delen, komt aan de orde hoe motorische controle van spraak zich ontwikkelt. Een van de uitdagingen bij het begrijpen van spraakstoornissen in het algemeen is om stoornissen in de fonologie te onderscheiden van stoornissen in de motorische controle (zie Kent, 2000). Zo is er in de literatuur over spraakapraxie bijvoorbeeld aanhoudende discussie of deze spraakstoornis bekeken zou moeten worden op het fonologische niveau, op het motorische niveau, of op beide. In dit eerste artikel komt, met als uitgangspunt het DIVA model voor spraakproductie (Guenther, 1994; 1995), een aantal belangrijke componenten van spraakmotorische controle aan bod. Hierbij wordt in het bijzonder aandacht besteed aan de manier waarop auditieve feedback een rol speelt bij de ontwikkeling en het handhaven van spraakmotorische controle. In het tweede artikel verschuift de nadruk van theoretisch naar meer klinisch en worden experimentele data besproken met betrekking tot de wijze waarop de vroege spraakontwikkeling verloopt. Hierbij wordt uitgebreid de ontwikkeling van brabbelen bekeken, aangezien dit voor kinderen een cruciaal stadium in de ontwikkeling van spraakmotorische controle vormt. Wederom neemt de rol van auditieve feedback een belangrijke plaats in bij het bespreken van de onderzoeksgegevens.

Inleiding

Wanneer men literatuur over de ontwikkeling van spraak en taal bekijkt dan wordt vaak benadrukt hoe ingewikkeld het leren spreken is. Gezien deze complexiteit is het opvallend hoe snel kinderen spraak-, en taalvaardigheden verwerven door als het ware “alleen maar te luisteren” naar volwassen spraak. In dit literatuuroverzicht, dat bestaat uit twee artikelen, zal aan de orde komen hoe kinderen geleidelijk steeds meer

controle krijgen over hun spraakproductie. Spraakproductie weerspiegelt de rijping van sensorische, motorische en cognitieve systemen (Kent, 2000). Tijdens de spraakontwikkeling zijn naast rijpingsfactoren ook een aantal externe factoren betrokken bij de verwerving van spraakmotorische controle. In de context van spraakmotorische controle zijn we met name geïnteresseerd in de wijze waarop perceptuele factoren zoals auditieve en visuele feedback ingrijpen op het proces van spraakontwikkeling. Hierbij zal in het eerste artikel een theoretisch kader besproken worden waarin aan de hand van een model voor spraakmotorische controle belangrijke componenten van motorische controle en de invloed van auditieve feedback aan bod komen. Vervolgens zullen in het tweede artikel experimentele data besproken worden over de wijze waarop de spraakontwikkeling van kinderen verloopt, waarbij uitgebreid de ontwikkeling van brabbelen zal worden bekeken, aangezien dit voor kinderen een cruciaal stadium in de ontwikkeling van spraakmotorische controle vormt. Hierbij zal wederom de rol van auditieve feedback een belangrijke plaats innemen.

Na een overzicht van de vroege spraakontwikkeling van kinderen en de invloed van auditieve feedback hierop, wordt, eveneens in het tweede artikel, de overstap gemaakt naar de spraakperceptie waarbij onderscheid wordt gemaakt naar auditieve perceptie en visuele perceptie. Op deze manier kan een indruk worden verkregen van factoren die naast auditieve feedback een invloed hebben op spraakproductie en de ontwikkeling hiervan.

Na zowel productie als perceptie besproken te hebben, wordt er tenslotte aandacht besteed aan de link tussen spraakproductie en -perceptie die tot uiting komt in imitatie. Hierbij zal met name ingegaan worden op de manier waarop imitatie een voor spraakmotorische controle zeer belangrijke link vormt tussen spraakperceptie en -productie.

Spraakmotorische controle

Voordat we een beeld schetsen van de wijze waarop spraakmotorische controle zich ontwikkelt, is het zinvol om te kijken naar wat onder spraakmotorische controle verstaan wordt. Kent (2000) geeft in zijn overzicht van motorische spraakstoornissen de volgende omschrijving van spraakmotorische controle (*speech motor control*): “spraakmotorische controle verwijst naar de systemen en strategieën die de productie van spraak controleren, inclusief de planning en voorbereiding van bewegingen (soms *motor programming* genoemd) en de uitvoering van bewegingsplannen die resulteren in spiersamentrekkingen en structurele verplaatsingen” (p 391 cursief aangebracht door deze auteurs). Over het algemeen wordt aangenomen dat de input van het systeem voor spraakmotorische controle een fonologische representatie van taal is, meer specifiek een reeks van abstracte eenheden zoals fonemen. De output van spraakmotorische controle is een serie articulatorische bewegingen die de geplande linguïstische boodschap overbrengen in de vorm van een akoestisch signaal dat door een luisteraar geïnterpreteerd kan worden. De processen van spraakmotorische con-

trole liggen dus tussen die van het formuleren van een talige boodschap en het realiseren van het akoestische signaal waardoor de boodschap van de spreker uitgezonden wordt.

Traditioneel wordt spraakmotorische controle onderscheiden van fonologische codering. Echter in enkele recente fonologische theorieën is er een vervaging van de grenzen tussen fonologische representaties en motorische functies. Toch is het onderscheid tussen fonologische en motorische processen zowel relevant voor het verklaren van spraakmotorische controle tijdens de normale spraakontwikkeling als voor het verklaren van verschijnselen van pathologische spraakontwikkeling. Zo is één van de uitdagingen bij het begrijpen van spraakstoornissen in het algemeen om stoornissen in de fonologie te onderscheiden van stoornissen in de motorische controle.

Gerelateerd aan deze kwestie rijst de vraag: wat is/zijn de belangrijkste variabele(n) die bij spraakproductie door het zenuwstelsel gecontroleerd dient/dienen te worden? Verschillende theorieën betreffende spraakmotorische controle trachten deze vraag te beantwoorden. Er zijn hierbij verschillende mogelijke variabelen voorgesteld, waaronder akoestische targets, aërodynamische waarden, posities van afzonderlijke bij articulatie betrokken structuren, algehele configuratie van het aanzetstuk en snelheden van spiersamentrekkingen. Hoewel veel studies zich op één specifieke variabele richten, is het mogelijk dat uiteindelijk meer dan één variabele nodig is om motorische controle van spraak te begrijpen. Aansluitend hierbij komt de robuustheid van spraak misschien gedeeltelijk voort uit de flexibiliteit van een spreker om tussen deze variabelen te kiezen waardoor het mogelijk wordt gemaakt om spraakproductie onder uiteenlopende omstandigheden onder controle te houden (Kent, 2000).

Kent (2000) noemt in zijn overzichtsartikel nog een andere manier om het bovengenoemde '*controlled variable probleem*' uit te drukken en dat is door te vragen welk soort 'target reference frame' door het neurale systeem gebruikt wordt om de articulatie van spraak te besturen. Kent (2000) bespreekt in dit kader twee belangrijke alternatieven namelijk: 1) Eén soort reference frame is een zogenaamd 'gestural target' dat gebaseerd is op de mate van constrictie op verschillende plaatsen in het aanzetstuk (Saltzman & Munhall, 1989). 2) Het andere reference frame betreft een auditief of akoestisch target dat gebruikt wordt om de articulatorische beweging af te leiden (Guenther, 1995; Guenther, Hampson & Johnson, 1998; Perkell et al., 1997). Er bestaat veel discussie over de aard van de target reference frames die door het systeem voor spraakmotorische controle gebruikt worden om de reeksen van fonemen (de input van het systeem) om te zetten naar continue articulatorische bewegingen voor het produceren van de bijbehorende spraakklanken. Hieronder zullen de twee belangrijke standpunten rondom de mogelijke targets voor articulatorische bewegingen tegen elkaar afgezet worden. Gezien de aanknopingspunten die een theorie met als uitgangspunt auditieve targets biedt voor het bespreken van auditieve feedback bij spraakmotorische controle zal deze uitgebreider besproken worden dan de theorie over gestural targets.

Constriction-based of auditief target?

De vraag bij deze uiteenzetting betreft dus: wanneer de articulatoren worden bewogen om een spraakklank te produceren betreft het invariante doel van het spraakmotorische controlesysteem dan een auditief doel (dat wil zeggen een doel gedefinieerd door parameters van het akoestische signaal zoals overgebracht door het auditieve systeem; (o.a. Guenther, 1995)), of dient het een meer kinematisch doel, zoals een target van constricties van het aanzetstuk als bewegingstarget? Met andere woorden: worden de bewegingstrajecten van de articulatoren gepland binnen een auditief reference frame of worden zij gepland binnen een reference frame dat nauwer gerelateerd is aan de vorm van het aanzetstuk? Zoals hierboven reeds vernoemd, is volgens Saltzman en Munhall (1989) het reference frame 'gestural' van aard, dat wil zeggen gebaseerd op de plaats en mate van constrictie van het aanzetstuk. Echter, volgens Guenther et al. (1998) is het reference frame voor de planning van spraakbewegingen een auditief doel, ook wel gelijkgesteld aan een akoestisch doel (Perkell et al., 2000).

Aanhangers van dit laatste standpunt betogen dat er een aantal observaties is dat door 'constriction based target' theorieën, zoals die van Saltzman en Munhall, moeilijk te verklaren is. Hoe verklaart een 'constriction based target' benadering bijvoorbeeld: (1) dat de meeste proefpersonen onder omstandigheden waarin bewegingen van hun lippen beperkt worden, in staat zijn om dezelfde klank te produceren als onder omstandigheden zonder dergelijke beperkingen en (2) dit gedeeltelijk doen door de plaats en mate van constrictie (constrictiepatronen) te veranderen (Callan et al., 2000; Guenther et al., 1998)? Het verschijnsel dat hetzelfde spraakdoel op meer dan een manier bereikt kan worden, wordt '*motor equivalence*' genoemd. Het is belangrijk dat een theorie over spraakmotorische controle een verklaring kan geven voor dit verschijnsel omdat kinderen in het dagelijks leven ook motor equivalence laten zien in hun vermogen om met verschillende vreemde objecten in hun mond spraak te produceren door verschillende articulatorische configuraties te gebruiken (Callan et al., 2000).

Een volgend voor constriction-based theorieën moeilijk te verklaren verschijnsel betreft de vraag hoe het centrale zenuwstelsel op grond van tactiele en proprioceptieve feedback de mate en plaats van constrictie in het aanzetstuk kan bepalen. Dit is problematisch daar deze 'mapping' van orosensorische feedback naar een representatie op grond van constricties verandert naarmate een kind groeit, omdat dan de fysieke kenmerken zoals de lengte van het aanzetstuk en de vormen van de articulatoren veranderen. Dergelijke veranderingen zullen van invloed zijn op het akoestische signaal dat wordt geproduceerd met een bepaalde verzameling motorische commando's. Een model voor een spraakmotorisch controlesysteem dient volgens Guenther (1995) een verklaring te geven voor de observatie dat het ondanks voortdurende veranderingen in de grootte, vorm en spierinnervatiepatroon van de articulatoren nog steeds mogelijk blijft om spraak te produceren. De veranderingen die tijdens de ontwikkeling optreden in de structuren betrokken bij spraakproductie hebben allemaal in variërende mate invloed op de akoestische resultaten en vereisen dat het kind constant leert om verschillende articulatorische configuraties te gebruiken om dezelfde spraakdoelen te

bereiken (Callan et al., 2000). Het spraakproductiesysteem moet dus een ‘*mapping*’ vestigen die in staat is om de articulatoren op een dusdanige manier te bewegen dat geleerde targets voor het produceren van fonemen bereikt worden, rekening houdend met de huidige bouw van het systeem die tijdens de ontwikkeling van het kind continu verandert. Bovenstaande gegevens lijken onverenigbaar met theorieën die uitgaan van invariante constriction-based targets (Callan et al., 2000).

Tevens is het niet duidelijk hoe tijdens de ontwikkeling op grond van tactiele en proprioceptieve feedback met betrekking tot de plaats en mate van constrictie een mapping tot stand kan komen die de beweging van de articulatoren stuurt om invariante constriction targets (of gestural targets) te bereiken. Precieze feedback over vorm van het aanzetstuk in termen van plaats en mate van constricties is over het algemeen niet direct beschikbaar voor het centrale zenuwstelsel (Guenther et al., 1998). Weliswaar kan informatie over plaats en mate van constrictie worden afgeleid uit beschikbare sensorische informatie, maar de beschikbare sensorische informatie is in zijn ruwe vorm niet georganiseerd in een ‘reference frame’ dat geschikt is voor de planning van spraakbewegingen. Dit zou vereisen dat een geleerde neurale mapping tussen de sensorische representaties en een neurale representatie in een constriction reference frame gebruikt wordt voor planning van bewegingen. Het leren van een dergelijke mapping wordt erg moeilijk, zo niet onmogelijk gemaakt door het gebrek aan een geschikt leersignaal en het gegeven dat er meerdere constricties mogelijk zijn voor het produceren van hetzelfde foneem.

Indien men auditieve feedback als trainingssignaal beschouwt dan is beter te verklaren hoe een mapping geleerd wordt die de beweging van de articulatoren stuurt om de targets voor spraakproductie te bereiken (Callan et al., 2000; Guenther et al., 1998). Dit idee wordt verder uitgewerkt in auditief perceptuele target theorieën die stellen dat auditieve feedback van zelfgeproduceerde spraak kan dienen als een adaptief signaal dat de bewegingen van de articulatoren stuurt om auditieve targets te bereiken (Callan et al., 2000; Guenther et al., 1998; Perkell et al., 1997).

Het DIVA model van spraakproductie

Een voorbeeld van een spraakproductiemodel dat een auditieve target referentieruimte gebruikt bij de planning van spraakproductie en tevens in staat is om een verklaring te geven voor het leren van de mapping tussen het auditief perceptuele frame en het articulatorische frame is het DIVA model (Guenther, 1994; Guenther et al., 1998). Het model wordt DIVA genoemd naar de mapping van *Directions* (in orosensory space) *Into* Velocities of Articulators (Guenther, 1994; 1995) die door het model (een neurale netwerk model) geleerd wordt.

Het DIVA model verschaft een verklaring voor uiteenlopende spraakproductievervalsingen, waaronder de reeds beschreven motor equivalence, variatie in productie van fonemen vanwege veranderingen in spreek snelheid, coarticulatievalsingen (onder andere Guenther, 1995) en is in staat om te laten zien hoe het spraakproductiesysteem omgaat met ontwikkelingsgerelateerde veranderingen in morfologie van de articulatoren (Callan et al., 2000). Het valt buiten het kader van dit artikel om in

detail te bespreken hoe het model met deze onderwerpen omgaat, maar aan de hand van de belangrijkste concepten uit het model zal besproken worden op welke manier een auditieve-naar-articulatorische mapping tot stand komt en welke rol auditieve feedback daarbij speelt.

Reference frames

Spraakproductie is misschien de meest complexe motorische taak die door mensen uitgevoerd wordt. Naast de verbazingwekkende snelheid waarmee woorden en fonemen worden uitgesproken, vereist het produceren van woorden ingewikkelde interacties tussen informatie uit veel verschillende reference frames (Guenther, 1994). Deze betreffen onder andere fonetische, akoestische, somatosensorische, en motorische frames. Een *fonetisch frame* beschrijft de klanken die een spreker wenst te produceren, en de signalen die deze klankeenheden aan een luisteraar overbrengen bestaan binnen een *akoestisch frame*. Akoestische signalen vormen het eigenlijke medium waardoor spraak gecommuniceerd wordt; de taak van het spraakproductiemechanisme is dan het creëren van een gepaste set van akoestische signalen om de linguïstische eenheden van de spreker naar de luisteraars over te brengen. *Somatosensorische signalen* van tactiele en proprioceptieve receptoren verschaffen informatie over de vorm van het aanzetstuk, die sterk bepalend is voor de geproduceerde klanken. Ten slotte worden *motorische reference frames* gebruikt om de commando's door te geven aan de individuele articulatoren en spieren om de bewegingen die in spraak resulteren te produceren. Normale spraakproductie komt voort uit het automatische gebruik van nauwkeurig afgestelde interacties tussen deze zeer verschillende reference frames (Guenther, 1994, 1995).

De parameters die de interacties tussen de frames regelen kunnen vanwege verschillende redenen niet vastliggen bij de geboorte van het kind, namelijk (1) taalspecificiteit: de precieze aard van de mappings tussen akoestische doelen en articulatiebewegingen is afhankelijk van de taal die gesproken wordt; (2) de verandering van de fysieke kenmerken (lengte en vorm) van de articulatoren tijdens de ontwikkeling van het kind; (3) eventuele tijdelijke of permanente beschadiging van de articulatoren. Bovengenoemde veranderingen beïnvloeden namelijk het akoestische signaal dat wordt geproduceerd met een bepaalde verzameling motorische commando's (Guenther, 1995). Het handhaven van het vermogen om op de juiste manier belangrijke akoestische kenmerken te produceren vereist dus dat parameters die de mapping tussen fonetische, akoestische, orosensorische en motorische frames regelen veranderen met de tijd.

In het DIVA model worden de verschillende mappings tussen reference frames geleerd tijdens een brabbelfase. Na deze brabbelfase is het model in staat om willekeurige combinaties te produceren van de fonemen die tijdens brabbelen geleerd zijn en om reeksen van fonemen om te zetten in articulatorische bewegingen voor het produceren van deze fonemen (Guenther, 1994; 1995; Guenther et al., 1998; Perkell et al., 2001).

Mappings in het DIVA model

In het (oorspronkelijke) DIVA model worden drie hoofdniveaus van representatie gebruikt, te weten een fonetisch niveau (spraakklank), een orosensorisch (ofwel somatosensorisch) niveau en een articulatorisch niveau (Guenther, 1994). Tijdens de brabbelfase leert het DIVA model twee mappings tussen deze niveaus: (1) Een fonetische-naar-orosensorische mapping waarbij voor iedere klank acceptabele ranges van orosensorische variabelen wordt geleerd; (2) een orosensorische-naar-articulatorische mapping waarbij gewenste bewegingen in orosensorische ruimte worden gekoppeld aan motorische commando's voor de articulatoren (Bailly, 1997; Guenther, 1994; 1995).

Callan en collega's (2000) noemen bij de beschrijving van hun versie van het DIVA model niet twee, maar drie mappings die tijdens het brabbelp proces door het model geleerd worden. Deze drie mappings betreffen de articulatorische-naar-auditieve mapping, de orosensorische-naar-articulatorische mapping, en de fonemische-naar-auditieve mapping.

Het verwerven van spreekvaardigheden bestaat in het DIVA model en afgeleiden hiervan uit het vinden van geschikte parameters voor de mappings die tijdens de brabbelfase geleerd worden (Callan et al., 2000; Guenther, 1994, 1995). Tijdens het leerproces maakt het model alleen gebruik van informatie die voor kinderen beschikbaar is (Guenther, 1995).

Intern model

De hierboven genoemde mappings worden verworven en gehandhaafd door middel van auditieve feedback (Perkell et al., 2000, 2001). Aangezien verwerkingstijden bij het gebruik van directe auditieve feedback voor spraakproductie te lang zouden zijn en dus voor te grote feedbackvertragingen zouden zorgen, is het onwaarschijnlijk dat directe auditieve feedback gebruikt wordt voor de controle van individuele articulatorische bewegingen en correctie van fouten in de spraakproductie (Perkell et al., 1997, 2000). Deze feedbackvertraging zou namelijk de tijd omvatten die nodig is voor zowel de perceptuele verwerking van de akoestisch-fonetische eigenschappen van de klank, de tijd die nodig is voor het corrigeren van motorische commando's, alsook de vertraging tussen spieractivatie en het genereren van spierspanning bij het produceren van corrigerende bewegingen. Net zoals bij andere typen geoefende bewegingen, kunnen bij spraakproductie problemen met te lange verwerkingstijden bij gebruik van auditieve feedback voor 'closed-loop' correctie worden vermeden door het gebruik van een intern model waarin het verband is vastgelegd tussen articulatie en klankoutput zoals geleerd tijdens spraakverwerving (zie o.a. Guenther et al., 1998; Perkell et al., 1997, 2000).

Een intern model is een representatie in het centrale zenuwstelsel van de functionele eigenschappen van reflexmatige, spier-, biomechanische en akoestische componenten van het spraakproductiemechanisme (Perkell et al., 1997). De term intern model wordt op een aantal manieren gebruikt (zie Jordan, 1990 in Guenther et al., 1998) en het is waarschijnlijk dat het brein meer dan één soort intern model gebruikt

bij het plannen van spraakbewegingen. Het interne model van het DIVA model is een aangeleerd model waarin het verband is vastgelegd tussen articulatie en klankoutput, met andere woorden op welke manier akoestische signalen gerelateerd zijn aan configuraties van het aanzetstuk. Bij het programmeren van articulatiebewegingen voor het bereiken van auditieve doelen wordt dus een intern model gebruikt waarin de verbanden tussen articulatorische configuraties en hun akoestische gevolgen vastgelegd zijn (Perkell et al., 2000).

In de literatuur over motorische controle van bewegingen is het begrip intern model een zeer geaccepteerd construct. Een mapping zoals vastgelegd in een intern model werd door Jordan (1990) een '*forward model*' genoemd en dit wordt ook teruggevonden in het DIVA model waarin een intern model dat informatie over de posities van de articulatoren omzet in een auditief perceptueel reference frame geleerd wordt met behulp van auditieve feedback (Guenther et al., 1998). Huidige modellen die een '*forward modeling component*' bevatten, steunen bij het vestigen van een mapping op een leersignaal. Het leersignaal (auditief, orosensorisch of articulatorisch) verschaft aan het forward model de output die het zou moeten produceren. Later kan de output van het forward model in plaats van het leersignaal worden gebruikt voor het produceren van de juiste klanken.

In het geval van het DIVA model wordt tijdens brabbelen een forward model geleerd dat informatie over de posities van articulatoren omzet in een auditief-perceptueel reference frame. Dit leren vindt als volgt plaats: tijdens brabbelen leiden posities van de articulatoren, gestuurd door het systeem, tot een akoestisch signaal. De posities van de articulatoren (informatie hierover is beschikbaar door middel van orosensorische feedback) functioneren als input voor het forward model. Het leersignaal voor het forward model is de auditieve terugkoppeling van de eigen spraak (Guenther et al., 1998). Op deze wijze leert het forward model zijn output af te stemmen op het leersignaal. Na het leren kan het forward model in plaats van auditieve feedback worden gebruikt om de huidige toestand van het aanzetstuk in het plannings reference frame (dat wil zeggen het auditief perceptuele frame) aan te geven, om vervolgens te bepalen welke commando's naar de articulatoren gestuurd moeten worden om het auditief perceptuele target te bereiken. Dit betekent dat, indien het forward model eenmaal geleerd is, het kan werken zonder de aanwezigheid van auditieve feedback (Guenther et al., 1998).

Interne modellen zijn bijzonder belangrijk geweest in recente spraakproductietheorieën en ze kunnen waardevol zijn bij het begrijpen van motorische spraakstoornissen (Jones & Munhall, 2000; Kent, 2000). In de volgende paragraaf zal specifieker ingegaan worden op de rol van auditieve feedback in de ontwikkeling van dit interne model, hetgeen kan verklaren dat continue, directe auditieve feedback tijdens volwassen spraakproductie niet onmisbaar is.

Auditieve feedback

Zoals verschillende keren naar voren gekomen is, moet het spraakproductiesysteem tijdens de ontwikkeling een mapping vestigen die in staat is om de articulatoren op

een bepaalde manier te bewegen om geleerde (auditieve) targets te bereiken. Een intern model waarin bovengenoemde mapping is vastgelegd moet door een zich ontwikkelend kind geleerd worden. Er wordt vanuit gegaan dat dit gebeurt met behulp van auditieve, somatosensorische en visuele feedback (Perkell et al., 1997; 2000). Met name voor auditieve feedback is een belangrijke rol weggelegd bij het vestigen van een dergelijke mapping of intern model (Guenther et al., 1998; Perkell et al., 1997). Callan et al. (2000) toonden in hun studie aan dat auditieve feedback betreffende zelfgeproduceerde spraak voldoende is om een mapping te trainen tussen auditieve target ruimte en articulatorische ruimte. Dit geldt zelfs onder omstandigheden waarin de structuren betrokken bij spraakproductie aanzienlijke ontwikkelingsgerelateerde herstructurering ondergaan (Callan et al., 2000).

Het bovenstaande gaat ook op voor het DIVA model, waarin mappings tussen articulatie en auditieve gevolgen (het interne model) worden verworven en gehandhaafd door gebruik van auditieve feedback. Hierdoor speelt auditieve feedback in het DIVA model vanzelfsprekend een belangrijke rol in de planning van spraakbewegingen (zie onder andere Perkell et al., 2001). Zoals reeds kort is opgemerkt, zijn verwerkingstijden te lang wanneer directe auditieve feedback zou worden gebruikt voor het bereiken van auditieve doelen door de articulatoren. Daarom wordt bij planning van spraakbewegingen een intern model gebruikt van de auditief-articulatorische mapping. Hoewel interne modellen mogelijk de noodzaak van continue directe auditieve feedback voor de controle van spraakbewegingen verminderen, blijft auditieve feedback nog steeds een belangrijke rol spelen in de verwerving en handhaving ervan (Jones & Munhall, 2000; Perkell et al., 1997, 2000, 2001). Aanvankelijk is auditieve feedback nodig om het interne model tijdens brabbelen af te stellen en steeds bij te stellen, maar naarmate het kind ouder wordt, wordt het interne model steeds robuuster en nauwkeuriger en wordt auditieve feedback vooral gebruikt om de parameters van het interne model (ofwel de parameters van de auditief-articulatorische mapping) te handhaven (Guenther et al., 1994, 1995, 1998; Perkell et al., 2000).

Volgens Jones en Munhall (2000) heeft auditieve feedback in formele modellen van akoestisch-articulatorische mappings zoals het DIVA model een aantal mogelijke rollen: (1) Voor de ontwikkeling van spraakklanken bij kinderen en volwassenen en voor het leren van nieuwe samenstellingen van het aanzetstuk verschaft auditieve feedback de primaire informatie over het al dan niet bereiken van het gewenste spraaktarget; (2) Voor fijnmotorische controle wordt de klank van de stem van de spreker gebruikt in closed-loop controle van de articulatie; (3) Voor motorische planning en controle verschaft het stemgeluid een doorlopende calibratie van interne modellen van het spraakmotorische systeem.

Empirische evidentie laat zien dat kinderen om normale spraak te verwerven moeten kunnen horen (bijv. Oller & Eilers, 1988). Tevens is het horen van zelfgeproduceerde spraak belangrijk voor het leren van taal en het handhaven van accurate articulatie (Jones & Munhall, 2000). Een manier om inzicht te krijgen in de rol van auditieve feedback bij spraakproductie is het bestuderen van sprekers die een verandering in gehoorstatus ervaren. Perkell en collega's (Perkell et al., 1992; Perkell et al.,

1997, 2000, 2001) bestudeerden de rol van auditieve feedback bij twee groepen die een verandering in gehoorstatus doormaakten na normale spraak te hebben verworven, namelijk:

1. cochleaire implant gebruikers die, na met tenminste adequaat gehoor spraak te hebben verworven, als volwassenen ernstig doof werden en vervolgens na enige tijd weer een vorm van kunstmatig gehoor kregen, en
2. patiënten met neurofibromatose type 2 (NF-2), die hun gehoor verloren vanwege tumoren die op de gehoorzenuwen groeien.

Studies over de spraak van deze patiënten die een verandering in gehoor doormaakten verschaffen op grond van veranderingen in de spraakwaliteit bewijs voor twee rollen voor auditieve feedback bij volwassenen spraakproductie:

1. Het handhaven van de parameters van het interne model. Een intern model van het verband tussen klank en articulatie wordt verworven en vervolgens gehandhaafd m.b.v. auditieve feedback (Perkell et al., 1997, 2000, 2001).
2. Het verzekeren van adequate verstaanbaarheid door middel van het monitoren van de akoestische omstandigheden om snelle, situatieafhankelijke aanpassingen te maken in houdingsparameters onderliggend aan ondermeer gemiddeld klankniveau, snelheid en grondtoonfrequentie (F0) (Perkell et al., 1997, 2000).

Ondanks dat longitudinale studies van postlinguaal doof geworden individuen suggereren dat auditieve feedback informatie verschaft die nodig is voor de lange termijn handhaving van accurate spraak van volwassenen alsook voor lange termijn veranderingen in de spraakproductie van een individu (Jones & Munhall, 2003), blijven mensen die postlinguaal doof zijn geworden (en dus met een normaal gehoor spraak hebben verworven), tot vele jaren na aanvang van de doofheid goed verstaanbaar spreken (hetzij soms enigszins afwijkend). Deze observatie betreffende goede verstaanbaarheid toont de robuuste aard van het interne model dat postlinguaal dove sprekers van het spraakproductieproces hebben verworven toen zij konden horen (Cowie & Douglas-Cowie, 1992; Perkell et al., 1992, 1997, 2000, 2001). Er zijn echter grenzen aan de robuustheid van het interne model. Uit de resultaten van de beperkte steekproef van Perkell en collega's (1997, 2000) bestaande uit 2 cochleaire implant gebruikers, mag voorzichtig geconcludeerd worden dat als spraakverwerving met gehoor voor het begin van ernstig gehoorverlies niet genoeg voortgang heeft geboekt, het resulterende interne model niet robuust genoeg is om met de verwerving van een vorm van kunstmatig gehoor tijdens de volwassenheid makkelijk hersteld te worden (Perkell et al., 1997, 2001).

Er zijn veel aanwijzingen dat vloeiende spraak gecontroleerd wordt door het gebruik van een intern model, alsmede door online verwerking van auditieve feedback (Jones & Munhall, 2000; Perkell et al., 1997). Een voorbeeld hiervan is de controle van de frequentie van de grondtoon (F0) tijdens het spreken. Jones en Munhall (2000) onderzochten in hun studie bij gezonde proefpersonen het belang van het gebruik van auditieve feedback bij het vormen van een intern model voor de controle van frequentie van de grondtoon (F0) bij spraakproductie. De data wezen erop dat auditieve feedback gebruikt wordt om zowel online correcties uit te voeren als voor

lange termijn afstelling van de mapping tussen spraakbewegingen en de resulterende auditieve feedback. Vermoedelijk gebruiken sprekers ook kinesthetische feedback om bij deze controle te helpen. Bovendien suggereren de data dat een soort intern model of representatie een rol speelt in de lange termijn afstelling van toonhoogte tijdens spreken (Jones & Munhall, 2000).

Naast de besproken resultaten uit experimenten bij gezonde proefpersonen, bestaan er ook observaties van de spraakproductie van postlinguaal doven die deze manier van controle van aspecten van spraakproductie illustreren. Zoals reeds vermeld vertonen mensen met verworven postlinguale doofheid vaak een verslechtering van veel aspecten van spraakproductie. Deze verslechtering kan geleidelijk plaatsvinden of vrijwel onmiddellijk. Na gehoorverlies worden vaak problemen gezien die gerelateerd zijn aan intensiteit en toonhoogtecontrole, alsook aan intonatie, klemtoon en spreesnelheid. Deze afwijkingen in de controle van luidheid, toonhoogte en spreesnelheid verschijnen over het algemeen kort nadat het gehoor verloren is gegaan (Jones & Munhall, 2003). Echter, (pas) na lange tijd van doofheid wordt ook variatie in de productie van klinkers en medeklinkers waargenomen (Cowie & Douglas-Cowie, 1992 in Perkell et al., 1997; Perkell et al., 1992). De bevinding dat precisie van klinker- en medeklinkerproductie tot een relatief lange tijd na het begin van de doofheid onveranderd blijft, ondersteunt het bestaan van een goedgevormde neurale mapping tussen het motorische systeem en de akoestische signalen voor spraaksegmenten (Jones & Munhall, 2000). Aan de andere kant impliceert de bevinding dat doofheid productieparameters zoals toonhoogte en intensiteit eerder aantast, dat de mechanismen betrokken bij suprasegmentele controle anders zijn dan die voor de controle van de productie van spraaksegmenten (Perkell et al., 1997). De controle van deze parameters zou gevoeliger kunnen zijn voor auditieve feedback (Jones & Munhall, 2000).

Auditieve feedback helpt niet alleen bij op peil houden van de spraakqualiteit bij gehoorverlies, maar de beschikbaarheid van auditieve feedback kan sprekers eveneens helpen om te compenseren voor structurele veranderingen in hun aanzetstuk zoals bijvoorbeeld een verlengstuk aangebracht aan de voortanden (Jones & Munhall, 2003). Auditieve feedback wordt dan door het spraakmotorische controlesysteem gebruikt om een onderliggende representatie van de mapping tussen articulatoire gestures en de akoestische gevolgen ervan te veranderen. Overigens is het waarschijnlijk dat een spreker bij een dergelijke structurele verandering in het aanzetstuk gebruik maakt van een combinatie van tactiele en auditieve feedback.

Naast auditieve feedback spelen somatosensorische en proprioceptieve informatie ook een belangrijke rol om het spraakmotorische controlesysteem te helpen om na veranderingen in het aanzetstuk spraak te reorganiseren en aan te passen aan nieuwe spraakomstandigheden (Jones & Munhall, 2003). Inzicht in het belang van sensomotorische interacties bij de coördinatie van spraakbewegingen komt van onderzoek naar spraakmotorische controle dat zich richt op studies waarbij het spreken verstoord wordt (onder andere Abbs & Gracco, 1984; Gracco & Abbs, 1985; Shaiman & Gracco, 2002). In dit soort studies worden tijdens spreken onverwachte verstoringen van een enkele articulator, zoals de kaak of de onderlip, teweeggebracht. Hierbij

is gevonden dat een mechanische verstoring van één articulator (bijvoorbeeld het mechanisch omlaag trekken van de onderlip) resulteert in compensatoire reacties in zowel de verstoorde als de niet verstoorde articulatoren die betrokken zijn bij het produceren van de specifieke spraakklank (Abbs & Gracco, 1984; Gracco & Abbs, 1985). Deze snelle compensatoire reacties vonden plaats binnen 60 ms (dat wil zeggen in minder dan de minimale reactietijd), wat aangeeft dat tactiele en proprioceptieve feedback een actieve rol spelen tijdens spraak. Meer specifiek zouden zij een rol spelen in het minimaliseren van spraakfouten en helpen bij het verfijnen van de spraakmotorische output (Abbs & Gracco, 1983; Shaiman & Gracco, 2002). Aangezien we ons in dit artikel voornamelijk richten op de rol van auditieve feedback bij de motorische controle van spraak, zal hier niet dieper ingegaan worden op de rol van sensomotorische en tactiele feedback in spraakproductie.

Conclusie

In dit eerste artikel zijn theorieën handelend over de ontwikkeling van spraakmotorische controle belicht. Het DIVA model, een neurale netwerk model van de motorische controle van spraak (Guenther, 1994; 1995), nam hierbij een belangrijke plaats in. Aan de hand van dit DIVA model is een aantal componenten van spraakmotorische controle, alsmede de rol die auditieve feedback speelt, uitgewerkt. Voor de ontwikkeling van spraakmotorische controle is het noodzakelijk dat het spraakproductiesysteem een mapping vestigt die in staat is om de articulatoren op een dusdanige manier te bewegen dat geleerde (auditieve) doelen voor het produceren van fonemen bereikt worden (met andere woorden, dat er een auditieve-naar-articulatorische mapping tot stand komt). In het DIVA model is dit verband tussen articulatie en klankoutput vastgelegd in een intern model. Dit interne model betreft een aangeleerd model waarin vastgelegd ligt op welke manier akoestische signalen gerelateerd zijn aan configuraties van het aanzetstuk.

Auditieve feedback speelt een cruciale rol bij het leren en handhaven van een dergelijk intern model. Aanvankelijk is auditieve feedback nodig om het interne model tijdens brabbelen af te stellen en steeds bij te stellen, maar naarmate een kind ouder wordt, wordt het interne model steeds robuuster en nauwkeuriger en wordt auditieve feedback vooral gebruikt om de parameters van het interne model (ofwel de parameters van de auditief-articulatorische mapping) te handhaven. In overeenkomst hiermee, leerden empirische data van postlinguaal dove volwassenen dat auditieve feedback van cruciaal belang is voor het verwerven en handhaven van een goede spraakqualiteit.

Aangezien auditieve feedback volgens het DIVA model tijdens de brabbelfase noodzakelijk is voor het leren van het interne model en daarmee voor de ontwikkeling van spraakmotorische controle, zal in het tweede artikel de rol van auditieve feedback tijdens de vroege spraakontwikkeling, en met name tijdens de ontwikkeling van brabbelen, bekeken worden. Tevens zal in dat artikel aandacht worden besteed aan de vroege spraakperceptie en de link tussen vroege spraakperceptie en -productie.

Summary

In this review, consisting of two papers, issues in the development of speech motor control are addressed. One of the challenges in understanding speech disorders in general is to distinguish impairments of phonology from impairments of motor control per se (see Kent, 2000). There is for example continuing discussion in the literature as to whether a given speech disorder, such as apraxia of speech, should be understood at the phonologic level, the motoric level, or both. In this first paper the DIVA model of speech production (Guenther, 1994; 1995) is used to present a number of important components of speech motor control. In this overview, the focus is on the role of auditory feedback in the development and maintenance of speech motor control. While the first article provides a more theoretical overview, the second article focuses on clinical and experimental data on early speech development. In addition, the development of babbling is extensively discussed since this is a crucial stage in the development of speech motor control. Again, the role of auditory feedback receives special attention in the discussion of the research literature.

Referenties

- Abs, J. H., & Gracco, V. L. (1983). Sensorimotor actions in the control of multi-movement speech gestures. *Trends in Neurosciences*, 6, 391-395.
- Abs, J. H., & Gracco, V. L. (1984). Control of complex motor gestures: Orofacial muscle responses to load perturbations of the lip during speech. *Journal of Neurophysiology*, 51, 705-723.
- Bailly, G. (1997). Learning to speak. Sensori-motor control of speech movements. *Speech Communication*, 22, 251-267.
- Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6, 201-251.
- Callan, D. E., Kent, R. D., Guenther, F. H., & Vorperian, H. K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language and Hearing Research*, 43, 721-736.
- Cowie, R. I., & Douglas-Cowie, E. (1992). *Postlingually acquired deafness: Speech deterioration and the wider consequences*. New York: Mouton de Gruyter.
- Gracco, V. L., & Abs, J. H. (1985). Dynamic control of the perioral system during speech: Kinematic analysis of autogenic and nonautogenic sensorimotor processes. *Journal of Neurophysiology*, 54, 418-432.
- Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics*, 72, 43-53.
- Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, 102, 594-621.
- Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, 105, 611-633.
- Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America*, 108, 1246-1251.

- Jones, J. A., & Munhall, K. G. (2003). Learning to produce speech with an altered vocal tract: The role of auditory feedback. *Journal of the Acoustical Society of America*, 113, 532-543.
- Jordan, M. I. (1990). Motor learning and the degrees of freedom problem. In M. Jeannerod (Ed.), *Attention and performance XIII* (pp. 796-836). Hillsdale, NJ: Erlbaum.
- Kent, R. D. (2000). Research on speech motor control and its disorders: A review and perspective. *Journal of Communication Disorders*, 33, 391-428.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1-36.
- Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development*, 59, 441-449.
- Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Perrier, P., Vick, J., Wilhelms-Tricarico, R., & Zandipour, M. (2000). A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *Journal of Phonetics*, 28, 233-272.
- Perkell, J., Guenther, F., Lane, H., Matthies, Vick, J., & Zandipour, M. (2001). Planning and auditory feedback in speech production. *Proceedings of the 4th International Nijmegen Speech Motor Conference*, Nijmegen, The Netherlands, June 13-16, 2001.
- Perkell, J. S., Lane, H., Svirsky, M. A. & Webster, J. (1992). Speech of cochlear implant patients: A longitudinal study of vowel production. *Journal of the Acoustical Society of America*, 91, 2961-2979.
- Perkell, J. S., Matthies, M. L., Lane, H., Guenther, F. H., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, 22, 227-250.
- Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1, 333-382.
- Shaiman, S., & Gracco, V. L. (2000). Task-specific sensorimotor interactions in speech production. *Experimental Brain Research*, 146, 411-418.